Title

Pre-clinical development of a lentiviral vector expressing the anti-sickling beta AS3 globin for gene therapy for sickle-cell disease

Journal

HUMAN GENE THERAPY, 28(12)

ISSN

1043-0342

Authors

Poletti, V
Charrier, S
Urbinati, F
et al.

Publication Date

2017

Peer reviewed

Molecular Therapy
Methods & Clinical Development
Original Article

Pre-clinical Development of a Lentiviral Vector Expressing the Anti-sickling βAS3 Globin for Gene Therapy for Sickle Cell Disease

Valentina Poletti,[1,7] Fabrizia Urbinati,[2] Sabine Charrier,[1] Guillaume Corre,[1] Roger P. Hollis,[2] Beatriz Campo Fernandez,[2] Samia Martin,[1] Michael Rothe,[3] Axel Schambach,[3,4] Donald B. Kohn,[2] and Fulvio Mavilio[5,6]

[1]Genethon, Evry, France; [2]Department of Microbiology, Immunology and Molecular Genetics, UCLA, Los Angeles, CA, USA; [3]Institute of Experimental Hematology, Hannover Medical School, Hannover, Germany; [4]Division of Hematology/Oncology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA; [5]Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy; [6]Paris Descartes University, Imagine Institute, Paris, France

Sickle cell disease (SCD) is caused by a mutation (E6V) in the hemoglobin (Hb) β-chain that induces polymerization of Hb tetramers, red blood cell deformation, ischemia, anemia, and multiple organ damage. Gene therapy is a potential alternative to human leukocyte antigen (HLA)-matched allogeneic hematopoietic stem cell transplantation, available to a minority of patients. We developed a lentiviral vector expressing a β-globin carrying three anti-sickling mutations (T87Q, G16D, and E22A) inhibiting axial and lateral contacts in the HbS polymer, under the control of the β-globin promoter and a reduced version of the β-globin locus-control region. The vector (GLOBE-AS3) transduced 60%–80% of mobilized $CD34^+$ hematopoietic stem-progenitor cells (HSPCs) and drove βAS3-globin expression at potentially therapeutic levels in erythrocytes differentiated from transduced HSPCs from SCD patients. Transduced HSPCs were transplanted in NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG)-immunodeficient mice to analyze biodistribution, chimerism, and transduction efficiency in bone marrow (BM), spleen, thymus, and peripheral blood 12–14 weeks after transplantation. Vector integration site analysis, performed in pre-transplant HSPCs and post-transplant BM cells from individual mice, showed a normal lentiviral integration pattern and no evidence of clonal dominance. An *in vitro* immortalization (IVIM) assay showed the low genotoxic potential of GLOBE-AS3. This study enables a phase I/II clinical trial aimed at correcting the SCD phenotype in juvenile patients by transplantation of autologous hematopoietic stem cells (HSC) transduced by GLOBE-AS3.

INTRODUCTION

Sickle cell disease (SCD) is caused by a mutation in the sixth amino acid of the hemoglobin (Hb) β-chain (E6V), which causes polymerization of Hb tetramers upon deoxygenation. As a consequence, red blood cells (RBCs) lose flexibility and adopt the characteristic sickle shape in the capillary circulation, causing ischemia, stroke, multi-organ damage, severe pain, chronic hemolytic anemia, and reduced life expectancy.[1] SCD is endemic in Africa and frequent in the Western world: approximately 100,000 individuals in the USA and 10,000 in France are affected by the disease, with an incidence of 1:5,000 and 1:2,500, respectively. Current treatment includes regular blood transfusions and induction of fetal Hb (HbF) synthesis by hydroxyurea.[2,3] The only definitive therapy for SCD is allogeneic transplantation of hematopoietic stem cells (HSCs) from matched sibling donors, with a reported disease-free survival rate of >90% at 6 years after transplantation.[4] Transplants with matched unrelated or mismatched donors carry progressively higher, unacceptable risks for morbidity and mortality.[4] Given the large variability in clinical severity, the limited availability of suitable donors, and the toxicity associated with the procedure, HSC transplantation is not frequently performed on SCD patients, particularly in the adult age. Gene therapy, i.e., transplantation of autologous HSCs genetically corrected with a lentiviral vector (LV) expressing an anti-sickling globin chain, could be a therapeutic option with less treatment-related morbidity, theoretically available to all patients. Hematopoietic stem-progenitor cells (HSPCs) transduced by LVs have been used in clinical trials of gene therapy for immunodeficiencies, lysosomal storage disorders, and hemoglobinopathies, providing strong evidence of clinical efficacy in the absence of treatment-related adverse events.[5]

Allogeneic HSC transplantation provides predictions about the minimal level of corrected HSCs and anti-sickling Hb synthesis necessary to achieve clinical benefit in SCD patients. Stable mixed chimerism with donor HSC levels as low as 10%–30% leads to significant clinical improvement, because of the selective survival of normal, donor-derived RBCs with respect to HbS-containing RBCs.[6–8] On the other hand, SCD patients carrying a hereditary persistence of HbF (HPFH)

Received 9 August 2018; accepted 29 October 2018;
https://doi.org/10.1016/j.omtm.2018.10.014.

[7]Present Address: Gene Therapy Program, Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Harvard Medical School, Boston, MA, USA.

Correspondence: Fulvio Mavilio, PhD, Department of Life Sciences, University of Modena and Reggio Emilia, Via Campi 287, 41125 Modena, Italy.
E-mail: fulvio.mavilio@unimore.it

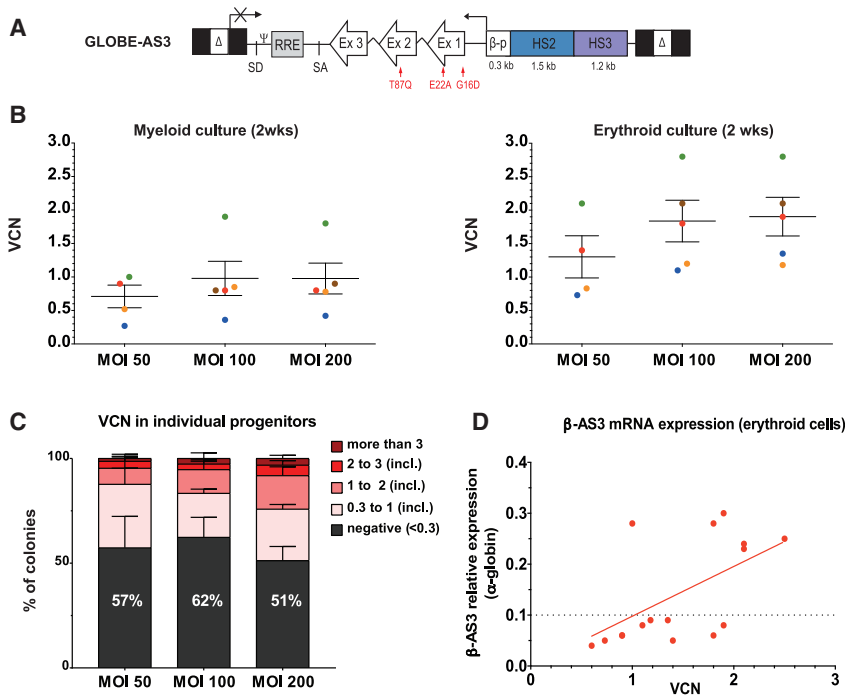

Figure 1. Transduction of Human Mobilized CD34⁺ HSPCs by the GLOBE-AS3 Vector

(A) Scheme of the GLOBE-AS3 lentiviral vector: the human βAS3-globin gene is under transcriptional control of a short (264-bp) β-globin promoter (β-p) and an HS2+HS3 mini LCR in anti-sense orientation in the CCL-SIN-18 LV backbone. Ex 1, 2, and 3 indicate the three β-globin exons. The three mutations are indicated in red. (B and C) Transduction efficiency of GLOBE-AS3 at increasing MOI in G-CSF-mobilized CD34⁺ HSPCs from five healthy donors, maintained for 2 weeks in liquid myeloid (B, left panel) or erythroid (B, right panel) culture, or cultured as individual progenitors (C) in semi-solid medium. (D) At day 14 of erythroid differentiation, cells were collected and the expression of βAS3 mRNA was determined by qRT-PCR as ratio between βAS3 mRNA and the endogenous α-globin mRNA, and correlated with VCN.

experience a much milder clinical course or are asymptomatic.[9,10] Together, these data suggest that engraftment of >20% of autologous HSCs producing an RBC progeny with >30% anti-sickling Hb levels could ameliorate the SCD pathology.

Pre-clinical and clinical studies have shown the potential of gene therapy in correcting the SCD phenotype. The LVs used in these studies express either the fetal γ-globin or mutant β-globins that interfere with axial and lateral contacts in the HbS polymer, thereby reducing HbS polymerization and sickling. The anti-sickling genes are expressed under the control of a β-globin promoter and combinations of elements from the β-globin locus control region (LCR) for maximal gene expression.[11] An LV expressing one such mutant, the βA^{T87Q}, is currently used in clinical trials for both β-thalassemia and SCD.[12,13] A second mutant, βAS3 globin, carries the T87Q and two additional mutations, G16D and E22A, which contribute to the anti-sickling activity and increase the affinity of the mutant for the α chain.[14] βAS3 globin, expressed in the context of an LV, corrects a murine model of SCD[15] and reduces the level of HbS and RBC sickling at potentially therapeutic levels when transferred in bone marrow (BM)-derived human CD34⁺ HSPCs from SCD patients.[16,17]

Here we present the characterization of an optimized vector (GLOBE-AS3) expressing the human β-AS3 gene under the control of a short human β-globin promoter and a reduced version (HS2+HS3) of the β-globin LCR. This vector design allows high transgene expression from a minimal size transgene, essential for high-titer vector production. As part of the pre-clinical validation of the vector, we analyzed the *in vitro* correction of the sickle phenotype in SCD patients' cells, as well as engraftment, biodistribution, and *in vivo* genotoxicity of transduced human HSPCs from healthy donors after xenotransplantation in an NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG) mouse model. A vector integration analysis was carried out before and after transplantation, to analyze the clonal dynamics of transduced cells *in vivo*, and positive or negative selection of cells harboring integration events in specific genes or classes of genes. This study has been designed to enable a multicenter phase I/II clinical trial aimed at establishing the safety and efficacy of HSPCs transduced with GLOBE-AS3 for gene therapy for SCD.

RESULTS

Design of a SIN LV for SCD Gene Therapy

A self-inactivating (SIN) LV (GLOBE-AS3; Figure 1A) was built by cloning the human βAS3-globin gene under the control of a short (264-bp) β-globin promoter and a HS2+HS3 mini LCR in anti-sense orientation in the CCL-SIN-18 LV backbone. The vector was pseudotyped with vesicular stomatitis virus glycoprotein (VSV-G) and produced by transient transfection in 293T cells with an average infectious titer of 3.5×10^9 transducing units (TU)/mL. Transduction efficiency was tested on CD34⁺ HSPCs mobilized with granulocyte-colony stimulating factor (G-CSF) in the peripheral blood (PB) of five healthy donors. HSPCs were transduced at increasing MOI (50, 100, and 200) and maintained for 2 weeks in liquid culture supporting myeloid or erythroid differentiation, or cultured as individual progenitors (colony-forming cells [CFCs]) in semi-solid medium. The average vector copy number (VCN) in the myeloid bulk cultures of transduced CD34⁺ cells was 0.7 ± 0.2, 0.9 ± 0.3, and 0.9 ± 0.2 (mean ± SEM), respectively, whereas it reached 1.3 ± 0.3, 1.8 ± 0.3, and 1.9 ± 0.3 in erythroid cultures (Figure 1B). The % of transduced individual CFCs (myeloid + erythroid + mixed colonies) was 49% at the highest MOI (Figure 1C). At day 14 of erythroid differentiation, ~80% of cells were positive for the erythroid markers CD36 and CD235a (data not shown). Cells were collected, total RNA was extracted, and the expression of βAS3 mRNA determined by qRT-PCR as a ratio between βAS3

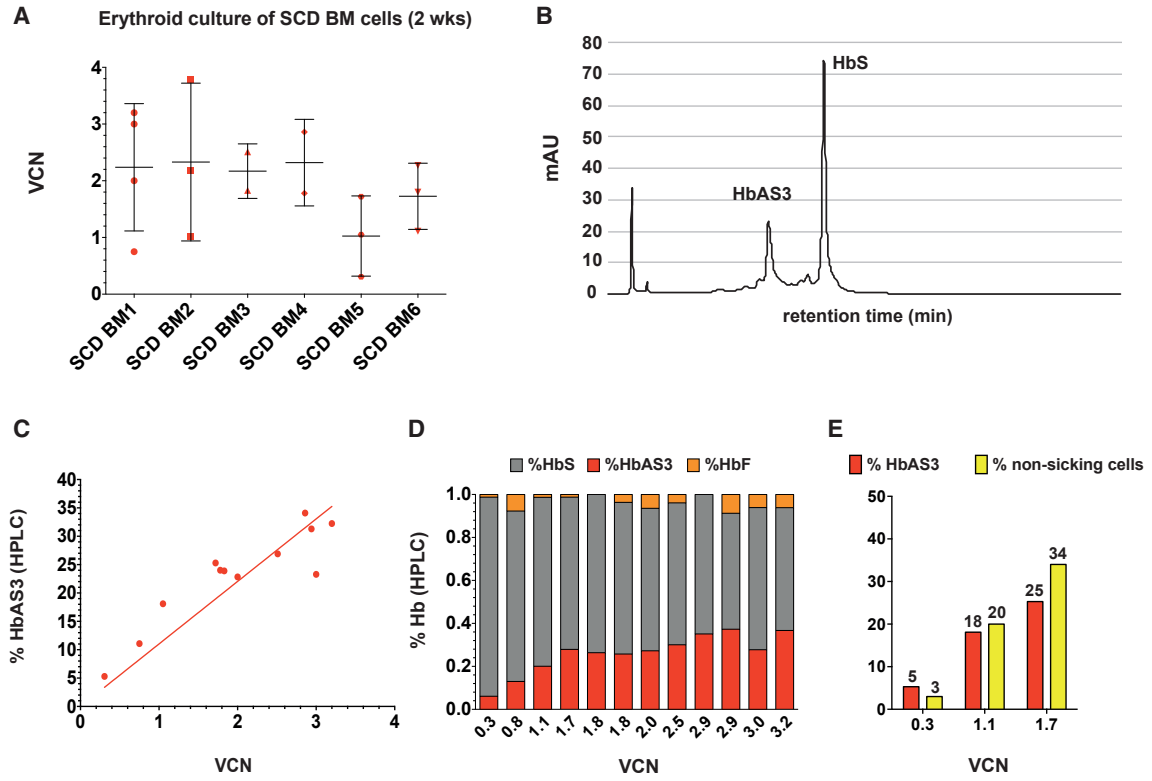

Figure 2. Correction of the Sickling Cell Phenotype *In Vitro* in Erythrocytes Expressing βAS3 Globin

(A) Average VCN in CD34$^+$ HSPCs from the BM of seven different SCD patients after transduction with GLOBE-AS3 at different MOIs (25–500), measured 2 weeks after transduction. (B) Quantification of HbAS3 tetramers by HPLC in a red blood cell (RBC) lysate. (C) Correlation between VCN and HbAS3 synthesis. Extrapolation of the correlation curve (R = 0.8305) estimates an output of 11 ng of HbAS3 per vector copy per cell. (D) The histogram shows the relative proportion of AS3, sickling (S), and fetal (F) hemoglobins in erythrocytes differentiated from CD34$^+$ cells with increasing VCN. (E) *In vitro* anti-sickling assay in erythrocytes differentiated in culture from BM CD34$^+$ cells from a representative SCD donor. RBCs derived from cells transduced with GLOBE-AS3 showed a percentage of phenotypically corrected, non-sickled forms proportional to the VCN.

mRNA and the endogenous α-globin mRNA (HbA). Relative βAS3 expression was correlated to the VCN, achieving between 4% and 30% of the α-globin expression with VCN values ranging from 0.6 to 2.5 (Figure 1D).

Expression of βAS3 Globin in RBCs from SCD Patients and Correction of the Sickle Cell Phenotype *In Vitro*

CD34$^+$ HSPCs were obtained from the BM of seven different SCD patients, transduced with GLOBE-AS3 at different MOIs (25–500), and maintained in erythroid cultures for 3 weeks. The VCN, measured 2 weeks after transduction, ranged between 0.3 and 3.8, with a substantial variability in cells from different donors transduced at similar MOI (Figure 2A), indicating a significant donor-dependent variability in transduction efficiency. After 3 weeks of culture, cells were harvested and the amount of vector-driven HbAS3 protein quantified by high-pressure liquid chromatography (HPLC) in RBC lysates (Figure 2B). A high correlation was observed between VCN and HbAS3 protein production (R = 0.8305; Figure 2C), with an estimated output of 11 ng of HbAS3 per integrated vector copy.

The effect of the synthesis of βAS3 globin on the SCD phenotype was tested by an *in vitro* anti-sickling assay in erythrocytes differentiated in culture from BM CD34$^+$ cells from one SCD donor. CD34$^+$ cells were transduced at MOIs of 45, 150, and 450 and cultured for 3 weeks in erythroid differentiation medium to obtain hemoglobinized, enucleated RBCs. Cells were harvested and incubated in sealed chambers with sodium metabisulfite to induce sickling as previously described.[16] Cell morphology was then examined under a phase-contrast microscope. RBCs derived from cells transduced with GLOBE-AS3 showed a higher percentage of phenotypically corrected, non-sickled forms compared with RBCs derived from mock-transduced cells from the same donor (Figure 2D). The percentage of phenotypically corrected cells correlated with VCN, reaching a maximum of 34.0% at a VCN of 1.7 (Figure 2D).

Transplantation of Human G-CSF-Mobilized CD34$^+$ Cells Transduced with LV-βAS3 in NSG Mice

CD34$^+$ HSPCs were mobilized by G-CSF from three healthy donors, pre-activated overnight with a cytokine cocktail, and either mock-transduced or transduced by two rounds of infection at MOI 100

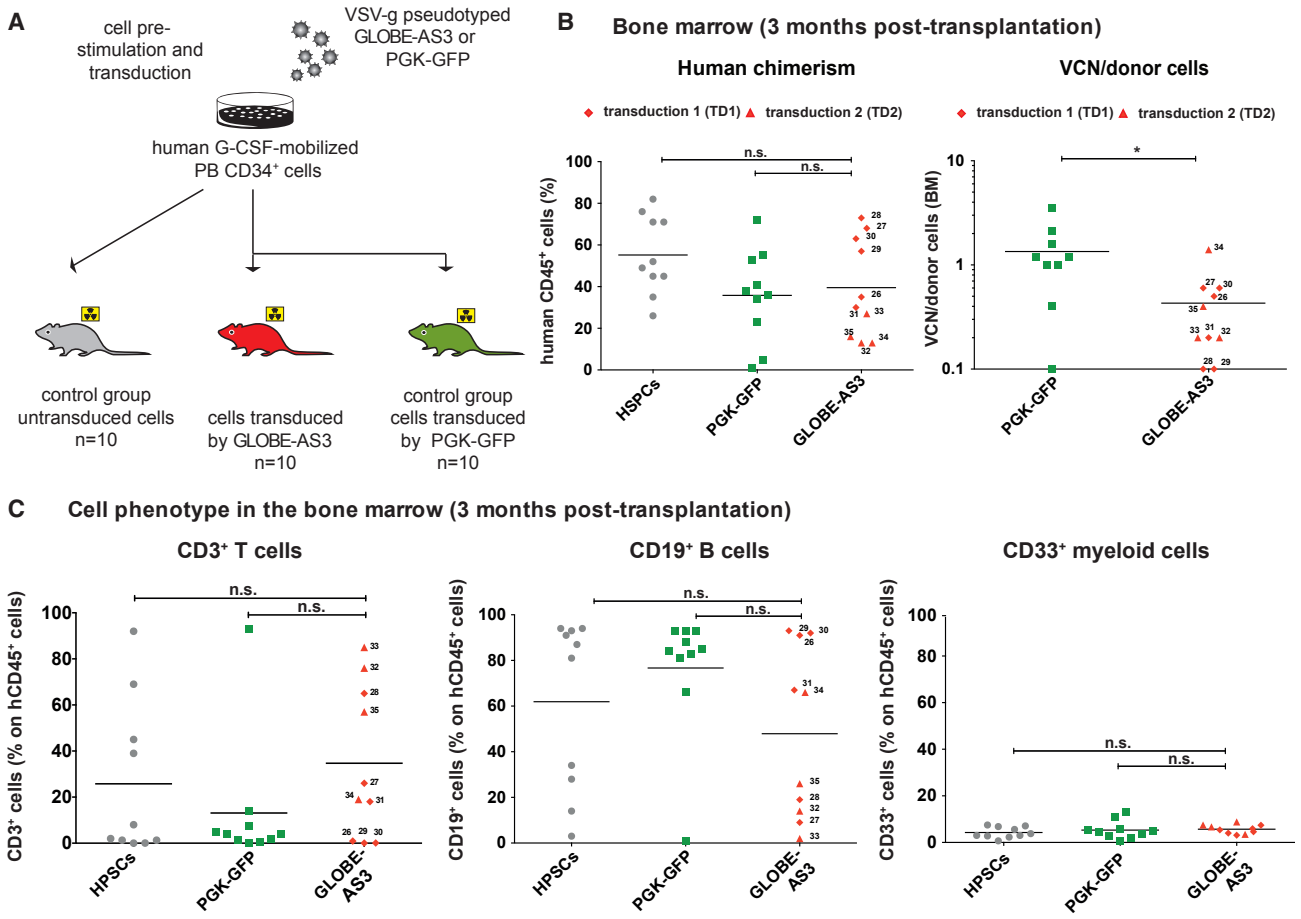

Figure 3. Transplantation of G-CSF-Mobilized CD34⁺ Cells Transduced with LV-βAS3 in NSG Mice

(A) Scheme of the *in vivo* study. CD34⁺ HSPCs from three healthy donors were mock-transduced or transduced with GLOBE-AS3, or with a control vector expressing GFP and transplanted in sub-lethally irradiated female NSG recipient mice. (B) Human cell chimerism (left panel) and average VCN in the donor cell fraction (right panel) were evaluated in the BM of mice 12 weeks after transplantation. Animals transplanted with cells coming from two independent transductions with GLOBE-AS3 (TD1 and TD2) are identified by red diamonds or triangles, respectively, and individual mice are identified by numbers. (C) The presence of human T (left panel), B (center panel), and myeloid (right panel) cells was evaluated by flow cytometry after immunostaining for human CD45⁺/CD3⁺, CD45⁺/CD19⁺, and CD45⁺/CD33⁺ cells, respectively. Bars indicate mean ± SEM. Statistical significance was calculated by a Mann-Whitney test: *$p < 0.05$. n.s., not significant.

with GLOBE-AS3 or with a control vector expressing GFP from the human phosphoglycerate kinase promoter (PGK-GFP). We performed two independent transductions, the first with CD34⁺ cells from one donor (TD1) and the second with cells pooled from two different donors (TD2). An aliquot of cells transduced with GLOBE-AS3 was maintained in liquid culture for a week for VCN evaluation and vector integration analysis or cultured as individual progenitors in semi-solid medium for 2 weeks. A VCN of 2.8 ± 0.2 and 4.7 ± 0.8 was obtained with GLOBE-AS3 and PGK-GFP, respectively, with 51% and 75% of transduced individual progenitors. Cells transduced with PGK-GFP were also analyzed for GFP expression by flow cytometry, resulting in 60.0% ± 9.0% GFP⁺ cells. After transduction, CD34⁺ cells were transplanted in sub-lethally irradiated, female NSG recipient mice (10 mice per group) by retro-orbital injection (~2×10^6 cells/mouse) (Figure 3A). Transplanted mice were maintained for 3 months and monitored weekly for health and body weight. One mouse that received untransduced cells and two mice that received cells transduced with GLOBE-AS3 were sacrificed at an early time point because of loss of weight due to the irradiation, with no significant difference among the three groups (χ^2 test, $p = 0.3$). White blood cell (WBC), RBC, and platelet counts were determined at sacrifice and showed no apparent difference among the three groups (Figure S1A).

At the end of the study, human cell chimerism was evaluated in the BM by flow cytometry after labeling with a human-specific anti-CD45 antibody. Engraftment of human cells was comparable in mice receiving cells transduced by the GLOBE-AS3 or by the control vector (Figure 3B). The average VCN/donor cell of the test group was ~0.4, showing efficient transduction of repopulating stem cells by the

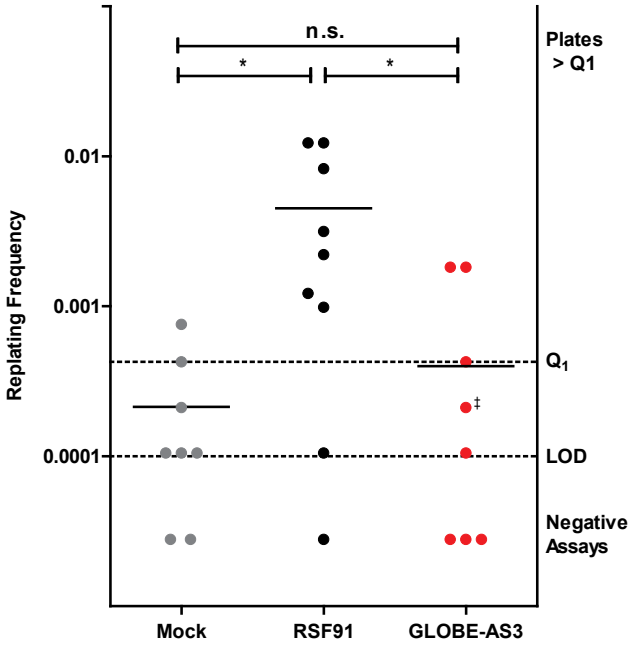

Figure 4. *In Vitro* Immortalization Assay

Proliferating clones were detected by MTT absorbance measurements in 96-well plates. Replating frequencies (RF) were calculated according to Poisson statistics. Each dot represents one assay. Values between the limit of detection (LOD) and the quantification threshold (Q1) were defined as background. All plates with clonal proliferation >Q1 were counted as positive. The negative assays below LOD were manually inserted into the graph. Bars indicate means. Statistical differences were calculated with a Fisher's exact test: *p < 0.05. Dagger (‡) indicates a sample with a very low VCN of 0.12. All other GLOBE-AS3 samples had an average VCN of 6.5 ± 0.5.

GLOBE-AS3 vector (Figure 3B). The average VCN was significantly lower than that obtained in cells transduced with the PGK-GFP vector, reflecting the lower transduction efficiency measured in pre-transplant cells. The presence of human T, B, and NK cells was evaluated by flow cytometry after immunostaining for human $CD45^+$/$CD3^+$, $CD45^+$/$CD19^+$, and $CD45^+$/$CD56^+$ cells, respectively. No significant difference in engraftment of immune cells was detected in the three groups of mice, with animals showing either predominant B cell or predominant T cell engraftment (Figure 3C; mice in the GLOBE-AS3 group are individually identified), a pattern reproduced also in PB and spleen (Figure S1B, top and middle panels). In animals where a thymus could be analyzed, this showed a high proportion of $CD4^+$/$CD8^+$ double-positive T cells, with no difference among the three groups (Figure S1B, bottom panel). As expected for this mouse model, engraftment of $CD45^+$/$CD33^+$ myelomonocytic cells was <10% in all hematopoietic organs (Figures 3C and S1B, top and middle panels), with no significant difference among the three groups. Overall, histopathological analysis of hematopoietic and non-hematopoietic organs revealed comparable size, cellularity, and organ architecture in mice receiving untransduced cells or cells transduced with either vector. No malignancy was observed in any of the groups (data not shown).

Analysis of the *In Vitro* Genotoxic Potential of the GLOBE-AS3 Vector

We tested the insertional mutagenesis potential of the GLOBE-AS3 vector in three independent *in vitro* immortalization (IVIM) assays.[18,19] As a positive control for vector-induced immortalization, we included a mutagenic gammaretroviral vector with intact LTRs. Non-transduced cells (Mock) were the negative control. Mutagenic, LTR-driven vectors like RSF91 contain strong enhancer-promoter sequences and can activate proto-oncogenes near the insertion sites.[20] In the IVIM assay, these mutants show a proliferation advantage when cells are seeded at very low density. While non-immortalized cells stop growing, insertional mutants can be quantified by their replating phenotype on a 96-well plate. We used a cumulative MOI of 400 in two rounds of transduction for GLOBE-AS3 and achieved a VCN of 6.5 ± 0.5 in 10 out of 11 transductions. One sample had a very low VCN of 0.12 for unknown reason. The positive control RSF91 reached VCN values of 8.8 ± 2.1. After transduction of 1×10^5 starting cells, cultures were expanded for 2 weeks and then re-plated at 100 cells/96-well plate. For RSF91, we observed insertional mutants in seven out of nine transductions with a mean replating frequency (RF) of 4.1×10^{-3}. In contrast, the GLOBE-AS3 vector showed a significantly lower incidence of positive plates (2 out of 11; $p = 0.022$, Fisher's exact test) and an over 10-fold lower RF in the IVIM assay, which was statistically not different from non-transduced Mock cells ($RF = 1.08 \times 10^{-4}$; $p = 0.05$, Kruskal-Wallis test with Dunn's correction) (Figure 4). Hence, GLOBE-AS3 showed a low risk for induction of insertional mutagenesis, even with multiple integrated proviral copies.

Analysis of the Integration Profile of GLOBE-AS3 in Human G-CSF-Mobilized $CD34^+$ HSPCs

The global integration characteristics of the GLOBE-AS3 vector were determined on an aliquot of transduced human HSPCs before transplantation. Integration sites (ISs) were recovered by linker-mediated PCR (LM-PCR) followed by high-throughput Illumina sequencing and mapped to the human genome (hg19 assembly) by a previously described, custom-designed bioinformatic pipeline[21] (Figure 5A). The aim of the analysis was to compare pre- and post-transplantation integration profiles to uncover differences in integration characteristics or frequency and function of the targeted gene pool that may suggest altered dynamics in the hematopoiesis of transplanted mice caused by positive or negative selection of cells harboring specific integration events.

The integration profile of GLOBE-AS3 in pre-transplant $CD34^+$ cells was analyzed in each transduction and in the merged TD1+TD2 dataset. Sequencing of the LM-PCR libraries generated 2.6–2.8 million reads/sample, corresponding to about 5.5 million reads in the merged TD1+TD2 dataset. Overall, a total of 41,979 ISs were retrieved from pre-transplant $CD34^+$ cells, with 13,459 and 28,721 unique ISs retrieved from TD1 and TD2 $CD34^+$ cells, respectively (Table 1). The GLOBE-AS3 vector showed the canonical LV genome-wide integration profile in $CD34^+$ cells, with a prevalence of intragenic, mostly

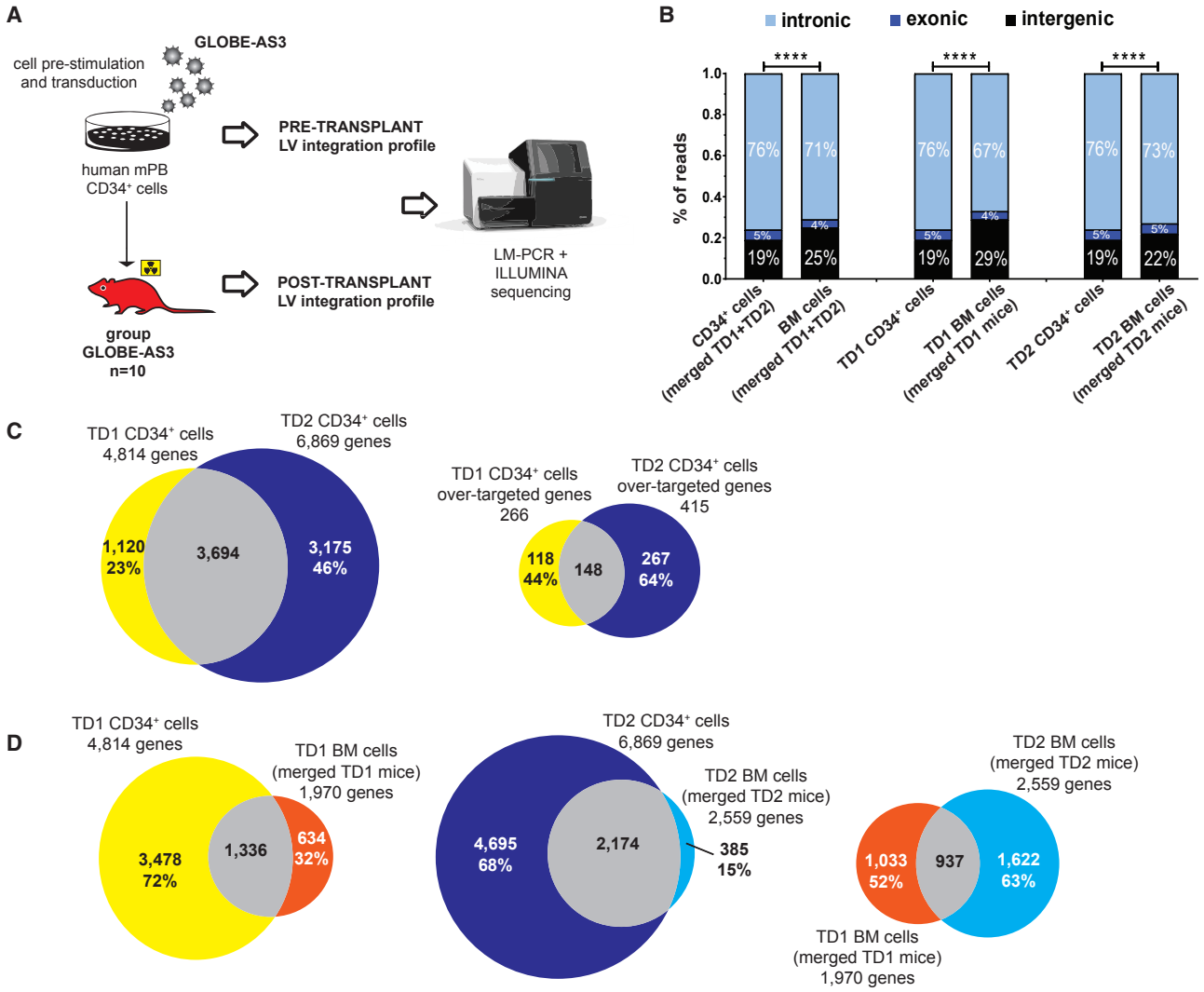

Figure 5. Integration Profile of GLOBE-AS3 in Human Mobilized CD34⁺ HSPCs

(A) The genome-wide integration pattern of the GLOBE-AS3 vector was determined on an aliquot of the transduced human HSPCs before transplantation in NSG mice. ISs were recovered by LM-PCR followed by high-throughput Illumina sequencing, and mapped to the human genome (hg19). (B) Distribution of ISs along the human genome in pre-transplant CD34⁺ cells (left bars) and the corresponding post-transplant BM cells (right bars). The proportion of intragenic (intronic or exonic) ISs was significantly decreased in post-transplant compared with pre-transplant cells (χ_2 test, ****$p < 0.0001$). (C) Venn diagram showing the proportion of targeted (left) and over-targeted (right) genes shared by TD1 (yellow) and TD2 (blue) samples in pre-transplant CD34⁺ cells. (D) Venn diagram showing the proportion of targeted genes shared by pre-transplant CD34⁺ cells and merged post-transplant BM cells in TD1 (left) and TD2 (middle) samples, and by the post-transplant TD1 and TD2 merged samples (right).

intronic ISs (81% in both TD1 and TD2; Figure 5B). Overall, the 41,979 ISs targeted 7,989 genes, corresponding to 28.5% of the annotated human genes. Most of the targeted genes were in common between TD1 and TD2, with the larger dataset (TD2) including the majority (77%) of the genes targeted in the smaller one (TD1) (Figure 5C).

We then analyzed the genes targeted at a statistically significant ($p < 0.001$) higher frequency with respect to an *in silico*-generated random set of integrations in the human genome, after Bonferroni correction for false discovery rate. By this analysis, 266 and 415 genes, respectively, were significantly over-targeted by GLOBE-AS3 in TD1 and TD2, respectively. Most of the genes over-targeted in TD1 (56%) were over-targeted also in TD2 (Figure 5C), and among the top-100 over-targeted genes (by IS frequency), 55 were in common between the TD1 and TD2 datasets (Table S1). The top-100 list contains most of the genes preferentially targeted by LV integration in HSPCs of patients treated by gene therapy for hematological and non-hematological disorders (in bold in Table S1). Among the targeted genes (merged TD1+TD2), several functional categories were enriched by DAVID analysis (Bonferroni corrected, $p \leq 0.001$), and included metabolism and transport of proteins, RNA and DNA, chromatin

Table 1. GLOBE-AS3 Integration Sites Retrieved from Pre-transplantation G-CSF-Mobilized PB CD34$^+$ HSPCs and BM Samples from NSG Mice 3 Months after Transplantation

Sample	Human Chimerism in BM (% huCD45)	VCN	No. of ISs	No. of Target Genes
Pre-transplant CD34$^+$ cells (merged TD1+TD2)	–	2.75	41,979	7,989
TD1 CD34$^+$ cells	NA	2.6	13,459	4,814
TD1 BM26	35	0.5	615	359
TD1 BM27	68	0.8	677	479
TD1 BM28	73	0.3	810	536
TD1 BM29	57	0.4	1,018	651
TD1 BM30	63	0.3	949	598
TD1 BM31	30	0.2	1,190	672
TD1 BM cells (merged)	54	0.4	3,822	1,970
TD2 CD34$^+$ cells	–	2.9	28,721	6,869
TD2 BM32	13	0.7	2,036	1,300
TD2 BM33	27	0.8	1,729	1,096
TD2 BM34	13	1.9	1,101	720
TD2 BM35	16	0.7	1,013	679
TD2 BM cells (merged)	17	1.0	5,332	2,559
BM cells (merged TD1+TD2)	–	–	8,870	3,595

huCD45, human CD45 antigen.

modifications, regulation of gene expression, and mitosis (the top-20 categories are listed in the Table S2).

Integration Profile of GLOBE-AS3 in Human NSG-Repopulating Cells

The integration profile of GLOBE-AS3 was then analyzed in the BM of each transplanted mouse in the TD1 and TD2 groups, as well as in the merged TD1+TD2 dataset, and the post-transplantation profiles compared with the corresponding pre-transplant CD34$^+$ datasets. Sequencing of the LM-PCR libraries generated 1–2.3 million reads/sample, corresponding to about 12 million reads in the merged TD1+TD2 dataset (Table 1). We retrieved between 615 and 2,036 ISs from individual BMs, suggesting a highly polyclonal repopulation in all transplanted mice (Table 1). On average, TD2 BMs yielded a higher number of ISs, in line with the higher average VCN observed in these animals and possibly reflecting a better transduction of the TD2 CD34$^+$ cells (Table 1).

The global post-transplant LV integration profile showed a prevalence of intragenic ISs (71% in TD1 and 78% in TD2), although their proportion was significantly decreased with respect to the corresponding pre-transplant CD34$^+$ samples (χ^2 test, $p < 0.0001$; Figure 5B), suggesting a negative *in vivo* selection of cells hosting intragenic integrations. The difference was significant also when considering only the integrations in exons (4.88% in the TD1+TD2 pre-transplant CD34$^+$ cells versus 4.44% in the corresponding post-transplant BM; χ^2 test, $p < 0.0001$; values rounded to 5% and 4% in Figure 5B). No selection for the orientation of the provirus with respect to the target gene was observed in the post- versus pre-transplant samples (Table S3). Most of the genes hosting a GLOBE-AS3 IS in the post-transplant BMs were targeted also in the corresponding pre-transplant CD34$^+$ dataset (68% in TD1 and 85% in TD2; Figure 5D), and no gene was targeted at a statistically higher frequency in BM cells (merged TD1+TD2) with respect to the corresponding pre-transplant samples. Almost half of the targeted genes (Figure 5D) and 20 out of the top-50 targeted genes were in common between the TD1 and TD2 BM samples (Table 2), and 23 (46%) and 29 (58%) of the top-50 targeted genes in TD1 and TD2, respectively, were listed in the top-100 over-targeted genes in the corresponding pre-transplant samples (second and fifth columns in Table 2). These data suggest no substantial positive or negative selection for cells carrying ISs in specific genes after transplantation. Again, most of the genes described as preferentially targeted by LV integration in patients treated by gene therapy are among the 50 top-targeted genes in BM cells (footnote "a" in Table 2).

A comprehensive functional analysis of the targeted genes in BM cells (merged TD1+TD2) performed by DAVID indicated no significant enrichment (Bonferroni-corrected $p \leq 0.01$) in functional categories with respect to the pre-transplant CD34$^+$ cells (merged TD1+TD2) except for the category "regulation of GTPase activity" (data not shown). Although no gene was targeted at a statistically higher frequency in TD1, TD2, and merged TD1+TD2 BM cells with respect to the corresponding pre-transplant samples, 34 genes were found enriched ($p < 0.05$ after Bonferroni correction) in individual BM samples (Table S4). Among them, only four genes were significantly enriched with $p \leq 0.001$, i.e., ORC4 in BM26, TECR and FHOD1 in BM28, and NDE1 in BM33 (in bold in Table S4), and none of them are present in cancer-related gene lists (see Materials and Methods).

Analysis of the Clonal Dynamics in Post-transplant BM Cells

The relative proportion of reads accounting for each individual IS in the pre- and post-transplant samples was determined and used as an indirect measure of the relative abundance of individual cell clones, considering an IS as a surrogate marker of clonality. All the ISs in pre-transplant CD34$^+$ cells (TD1, TD2, and merged TD1+TD2) and 88% of the ISs in the post-transplant BM cells (merged TD1+TD2) were represented by <1% of the total reads (Figure 6), indicating the largely polyclonal composition of these cell samples. In the BM of individual mice, some ISs had read counts >1% of the total reads of the sample (Table S5), possibly reflecting a higher abundance of some NSG-repopulating cell clones at the time of sample collection. In particular, three ISs showed read counts >10% in the BM of mice belonging to the TD1 group (Table S5). Overall, this group showed a less polyclonal BM repopulation and a consequent over-representation of individual cell clones. These data are in line with the lower transduction and lower VCN of TD1 CD34$^+$ cells with respect to TD2 cells (Table 1), and the lower IS recovery in the BM of individual mice (on average, 876 versus 1,470 ISs/mouse in TD1- versus TD2-transplanted mice; Figure 6).

Table 2. Top 50 Genes Targeted by GLOBE-AS3 in the BM of NSG Mice Transplanted with the TD1 and TD2 CD34$^+$ Cells, Ranked by Number of Independent ISs per Gene

TD1			TD2		
Gene Name	No. of ISs	Shared with TD1 CD34$^+$ Cells	Gene Name	No. of ISs	Shared with TD2 CD34$^+$ Cells
FCHSD2	16	X	**PACS1**	28	X
NBPF20	11		**KDM2A**	18	X
EYA3	9	X	**NF1**	17	
NF1	9	X	EIF4G3	14	X
EHMT1	8	X	PHACTR4	13	
GRB2	8	X	**FCHSD2**	12	X
PACS1	8	X	BRE	11	
ERC1	7		GATAD2B	11	X
GPATCH8	7	X	NBPF20	11	
KDM2A	7	X	WASF2	11	
NFATC3	7	X	**GPATCH8**	10	X
PHACTR4	7	X	MIR1268A	10	X
ACOX1	6	X	**ASH1L**	9	X
ASH1L	6	X	GRB2	9	X
C6orf10	6		NDE1	9	
FNBP1	6	X	NSD1	9	X
GATAD2B	6		PBX3	9	X
IP6K1	6	X	PPP6R2	9	X
RAB11FIP3	6	X	RAD51B	9	
RERE	6		THADA	9	
SARNP	6	X	CAPN1	8	X
STK4	6		EHMT1	8	X
VAV1	6	X	FBXL20	8	
ATF7	5	X	FNBP1	8	X
C11orf30	5		SMG6	8	X
EIF4G3	5	X	STAT3	8	
FHOD1	5		STAT5B	8	X
GON4L	5	X	STK4	8	
LRBA	5		TAOK1	8	
NAALADL2	5		**ACOX1**	7	X
NBPF10	5		BLM	7	X
NBPF9	5		CARD8	7	X
NFAT5	5		CLIC4	7	
ORC4	5		MGA	7	X
RPTOR	5		NBAS	7	
ST3GAL3	5	X	**NPLOC4**	7	X
TAOK1	5	X	SKAP1	7	
TECR	5		SMG1	7	X
TRPC4AP	5		SPATS2	7	X
DIP2B	4		UBR2	7	X
FOXJ3	4		CEP85	6	
FRY	4		FAM168A	6	
OS9	4		FLYWCH1	6	
POLE	4		NBPF9	6	X
RABGAP1L	4		NEAT1	6	
RBMS2	4		NFAT5	6	X
RPRD2	4		**NFATC3**	6	X
SBF2	4		RERE	6	
SDHC	4		TMEM57	6	
VPS45	4	X	TTBK2	6	X

For each gene, an X indicates that the gene was also targeted in the respective pre-transplant CD34$^+$ cell sample. Genes targeted in both TD1 and TD2 BM samples are highlighted in matching colors. Genes found preferentially targeted by LV integration also in gene therapy patients are indicated in bold.

In the TD2 group, only one BM sample had an IS with a read count higher than 5% of the total reads (BM35, chr1, position +153821982 in the GATAD2B gene; Figure 6), whereas all BM samples of the TD1 group showed ISs with read counts >5% of the total reads/sample (Figure 6). In particular, one IS in the TD1 BM27 sample accounted for 48% of the total reads, i.e., an intronic insertion in anti-sense orientation in the FOXP1 gene (chr3, +71173918; Figure 6), a Forkhead-box transcription factor involved in multiple cell differentiation pathways during development and differentiation. Neither this IS, nor any of the other ISs with read count >1% in individual BM samples, exceed 1% of the reads in the other BM samples of the same TD group, suggesting a random rather than a consistent, IS-dependent clonal amplification event.

Overall, 81% of the ISs accounting for >1% of the total reads/sample were intragenic and targeted 146 genes, which were not enriched for gene ontology (GO) categories or KEGG pathways by DAVID functional analysis and were not included in the list of genes previously described as enriched in targeting frequency in the BM of individual mice. These data again suggest a random rather than integration-driven clonal fluctuation in the BMs of the transplanted mice.

As expected, only a few of the ISs mapped in the BM samples were found in the corresponding pre-transplant sample (Figure 7; Table S6), indicating the exceedingly polyclonal composition of both TD1 and TD2 samples. Interestingly, some ISs were in common among different mice in each transduction group, suggesting that a certain number of NSG-repopulating cells expanded in the transduction culture and were able to repopulate more than one mouse (Figure 7; Table S6). Contamination between different BM samples was excluded by the DNA bar coding used in the LM-PCR and sequencing procedure (see Materials and Methods).

DISCUSSION

SCD is the most frequent monogenic disease worldwide, affecting millions of individuals.[22] Gene therapy, or the autologous transplantation of genetically corrected HSCs, is a potential alternative to allogeneic HSC transplantation, carrying lower transplant-related risks and being theoretically available to all patients. LV-transduced HSPCs have been used in clinical trials of gene therapy for immunodeficiencies and lysosomal storage disorders, providing strong evidence of safety and long-term efficacy in most cases.[5] Gene therapy for SCD faces additional challenges: although the size of the gene expression cassette has been reduced to a minimum in currently used vectors, LCR-β-globin vectors remain complex and have a lower transduction efficiency in HSCs compared with simpler vectors. The clinical history of allogeneic HSC transplantation provides some predictions about the minimal level of corrected HSCs necessary to achieve clinical benefit.[6–8] Recent gene therapy clinical trials showed, however, that the reduced output of a vector-borne transgene with respect to the natural β-globin allele significantly increases the minimal chimerism necessary to achieve clinical efficacy.[12,23] In addition, vector-derived β-globins must compete for α chains with the output of two endogenous β-globin genes in order to reduce sickling, requiring higher VCNs in the transduced cell products. SCD is therefore a more demanding context with respect to β-thalassemia, where substantial clinical benefit has been achieved with the current β-globin vectors.[13,24,25] Innovation in vector design, HSC procurement, cell manufacturing, and conditioning regimens appear necessary to achieve the efficacy required to cure SCD.

In an attempt to increase both transduction efficiency and anti-sickling activity of the therapeutic LV, we decided to use the βAS3 gene carrying three different anti-sickling mutations, T87Q, G16D, and E22A.[14] The mutations interfere with both axial and lateral contacts of β-chains within the HbS polymer and also increase affinity for the α chains, potentially providing a higher anti-sickling activity compared with the single T87Q

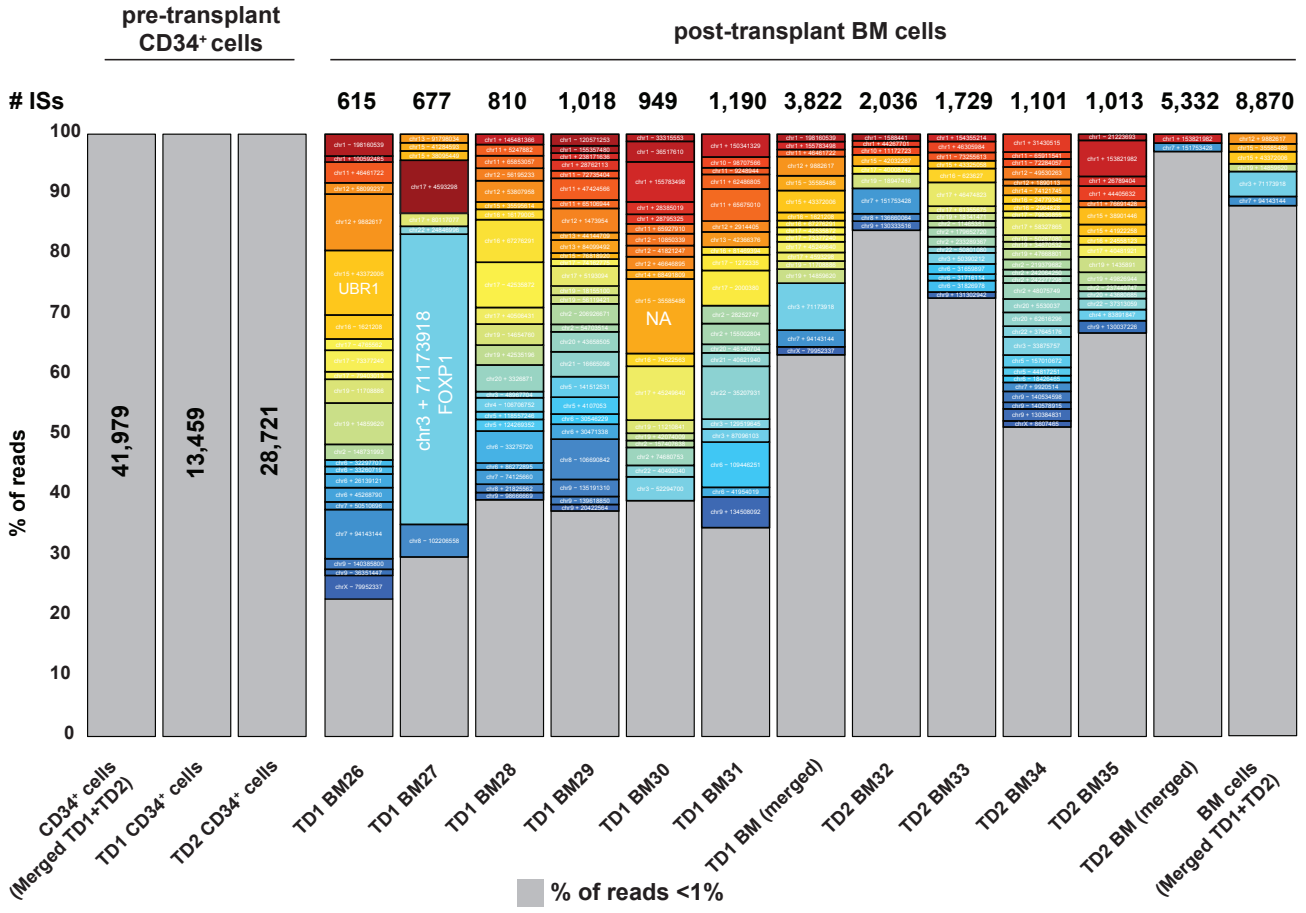

Figure 6. Clonal Dynamics in BM Cells of Transplanted NSG Mice

Relative proportion of reads accounting for each individual IS in pre- and post-transplant samples from individual NSG mice. The gray area in the bars indicates ISs accounting for <1% of the total in each sample. Individual ISs >1% are color-coded. Three ISs showed read counts >10% in the BM of mice 26, 27, and 30 in the TD1 group (see also Table S5), in line with lower transduction level and VCN in TD1 CD34$^+$ cells with respect to TD2.

mutant globin carried by the vector currently in clinical development.[12] The βAS3 gene was inserted in the GLOBE vector, which contains the HS2 and HS3 elements of the LCR and lacks HS4, which has a detrimental effect on LV titer and infectivity.[26] The enlarged size of the HS2 element increases its enhancing activity and compensates for the absence of HS4.[27] Overall, the GLOBE vector has a reduced size and a relatively higher titer and infectivity with respect to more complex vector containing HS4 and the β-globin 3′ enhancer, and shows unsurpassed clinical efficacy in transfusion-dependent β-thalassemia patients.[25]

The GLOBE-AS3 vector transduced mobilized CD34$^+$ HSPCs at high levels, as estimated by overall VCN in the transduced bulk culture and by analysis of individual clonogenic progenitors differentiated *in vitro*. HPLC analysis of globin chains in erythrocytes differentiated from transduced HSPCs obtained from the BM of SCD patients showed that GLOBE-AS3 drives βAS3-globin expression at potentially therapeutic levels (>30%[6–8]) and corrects to a significant extent the sickling phenotype in an *in vitro* assay. The level of phenotypic correction was linearly correlated with the average VCN in the bulk culture, indicating that two or more vector copies per genome are probably necessary to approach therapeutic efficacy, an estimate supported by the available clinical data.[12,23] A VCN of 2 was obtained in the majority of the SCD donors in the study in standard transduction conditions, an efficiency that can be further improved in a clinical transduction setting by using plerixafor-mobilized cells[28] and adding hormones or adjuvants to the transduction medium. Combined to the preliminary results coming from the β-thalassemia clinical trial,[25] these data indicate that GLOBE-AS3 is a promising vector for gene therapy of SCD.

Vector IS analysis carried out on transduced HSPCs showed canonical LV integration patterns, with the majority of integrations occurring within genes and the expected high frequency of targeting of a

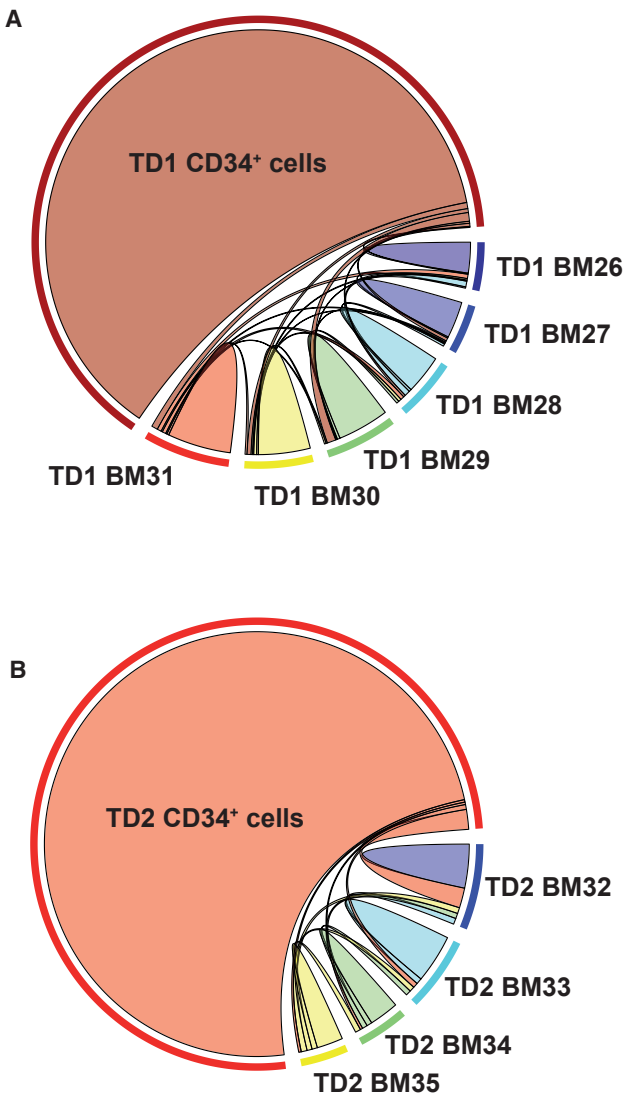

Figure 7. Clonal Sharing in BM Cells of Transplanted NSG Mice

Only a few ISs mapped in the post-transplant BM samples were identified also in the corresponding pre-transplant cells because of the highly polyclonal composition of both TD1 (A) and TD2 (B) samples. Interestingly, some ISs were shared by different mice in each transduction group, suggesting a number of NSG-repopulating cells expanded before transplantation and were able to repopulate more than one mouse.

defined sub-set of "hot" genes, as previously reported for other LVs in pre-clinical and clinical studies.[29] The list of over-targeted genes was consistent between two different transduction batches, showed no bias for particular functional categories of genes, and contained most of the genes previously reported as over-targeted in clinical studies of gene therapy for hematological and non-hematological disorders.[29–32] These results indicate that the GLOBE-AS3 vectors do not deviate from the known LV integration biases in HSPCs, which depend on the epigenetic and topological features of chromatin in this cell type.[29] An IVIM study showed low genotoxicity *in vitro* in a sensitive clonal cell proliferation assay, showing no departure from what was previously reported for clinical LVs by the same assays. This indicates that the combination of LCR-derived elements and β-globin promoter assembled in GLOBE-AS3 has very low gene activation potential in undifferentiated hematopoietic progenitors.

As part of the pre-clinical validation of the vector, we analyzed engraftment, biodistribution, and *in vivo* genotoxicity of transduced human HSPCs mobilized from healthy donors after xenotransplantation in an NSG mouse model. Vector integration analysis at the end of the study showed that BM repopulation was sustained by a largely polyclonal repertoire of HSCs, with no signs of vector-driven clonal dominance and no significant selection of cells carrying integration in specific genes, as evaluated by comparing pre-transplant and post-transplant dataset in individual animals. Unexpectedly, we found a few integrations in common among different mice transplanted with the same transduced cell batch, suggesting that at least some NSG-repopulating cells do expand in the pre-activation and transduction culture period without losing repopulation capacity. Alternatively, although less likely, these findings may indicate that LVs have a certain probability of integrating at exactly the same nucleotide in different repopulating cells. Interestingly, we observed a slight but significant reduction in intragenic integrations in NSG-repopulating cells with respect to pre-transplant HSPCs, indicating a negative *in vivo* selection of cells carrying integrations in exons and introns. Integrations in exons cause mono-allelic gene knockout, whereas the presence of LV proviruses in introns may induce post-transcriptional alterations of gene expression through the insertion of splicing and polyadenylation signals in primary transcripts, as previously reported.[33,34] These data suggest that some integration events may reduce cell fitness *in vivo* and be counter-selected when occurring in repopulating HSCs. Nevertheless, we observed no significant differences in the categories of genes targeted by LV integration before and after transplantation, indicating that if intragenic integrations cause reduction of cell fitness they do so in a relatively random fashion.

Overall, based on the prediction power of the *in vitro* and *in vivo* tests used in our study, the safety profile of the GLOBE-AS3 vector appears comparable with that of any other clinical LV. Combined with an accurate selection of the target cell population, state-of-the-art cell processing, and a patient conditioning regimen, the safety profile, transduction efficiency, and anti-sickling activity of GLOBE-AS3 pave the way to its clinical development in gene therapy for SCD.

MATERIALS AND METHODS

Vector Production and Titration

LVs pCCL_HS2+HS3_βp_βAS3 (GLOBE-AS3) and pCCL_PGK-GFP (PGK-GFP) were produced by co-transfection of HEK293T cells with CMV-GAG/POL, CMV-REV, and CMV-VSV-G plasmids in

225-cm^2 tissue culture flasks. After collection of the viral supernatant 24 and 48 hr post-transfection, the virus was concentrated by ultracentrifugation and resuspended in X-VIVO 20 medium (Lonza). For the IVIM and the animal study, GLOBE-AS3 was produced by transfection of HEK293T in cell factories (CF10), purified by ion-exchange chromatography, concentrated by tangential-flow filtration, and formulated in X-Vivo 20 medium, with a process identical to that used to produce clinical-grade vectors in good manufacturing practice (GMP) conditions. The viral titer was calculated by transduction of a standard human colorectal carcinoma cell line (HCT116) with serial dilution of the viral preparation, followed by VCN determination.

Human Cell Transduction and Transplantation into NSG Mice

$CD34^+$ cells were isolated from G-CSF-mobilized PB samples provided by the Centre Hospitalier Sud Francilien (Evry, France) or by the Institut Gustave Roussy (Paris, France), under full informed consent. $CD34^+$ cells were immunoselected with the CliniMACS CD34 reagent system (Miltenyi Biotec) following the manufacturer's instructions, pre-activated by overnight culture in X-vivo 20 (Lonza) containing 300 ng/mL human stem cell factor (SCF) (CellGenix), 300 ng/ml human Flt-3 ligand (CellGenix), 100 ng/mL human thrombopoietin (TPO) (CellGenix), and 20 ng/mL human interleukin-3 (hIL-3) (CellGenix). For the *in vitro* study, cells were transduced overnight with increasing MOI of LV-βAS3 and maintained in myeloid or erythroid differentiation media. Myeloid differentiation medium was formulated as X-vivo 20 medium containing 50 U/mL penicillin, 50 ng/mL streptomycin, 100 ng/mL human SCF (hSCF), 100 ng/mL hFlt-3 ligand, 100 ng/mL hTPO, and 20 ng/mL hIL-3. Erythroid differentiation medium was composed of StemSpan medium supplemented with 50 U/mL penicillin, 50 ng/mL streptomycin, 20 ng/mL hSCF, 1 U/mL human erythropoietin (hEPO), 5 ng/mL hIL-3, 2 μM dexamethasone, and 1 μM β-estradiol. Cells were maintained in differentiation medium for 12–14 days at a density of 0.3–0.5 $\times$ 10^6 cells/mL, with media changes every second or third day, as needed. At the end of culture, 0.5–1 $\times$ 10^6 cells were pelleted for DNA or RNA extraction.

For the animal study, cells from three different donors were used and transduced in two individual transductions, TD1 (MPB_16-06_W) and TD2 (pool of MPB_16-05_M and MPB_15-04_M). Cells underwent two hits of transduction (24 hr each) with the LV-βAS3 vector at an MOI of 100 in the same medium at the concentration of 1 $\times$ 10^6 cells/mL, in the presence of protamine sulfate (4 μg/mL; Sigma-Aldrich). After transduction, cells were washed and transplanted into sub-lethally irradiated NSG recipient mice. An aliquot of transduced cells was maintained in liquid culture in myeloid medium for 1 week, or grown as individual progenitors in semi-solid Methocult medium (H3434; STEMCELL Technologies) for 2 weeks. Individual colonies were counted under an inverted microscope after 1 week of culture, and collected for VCN evaluation at 2 weeks (50 colonies per condition).

Nine-week-old female NSG mice (Jackson Laboratory Stock No. 005557, supplied by Charles River) were hosted at the Genethon animal facility at CERFE (Evry, France) in accordance with all national and European ethical guidelines. Transduced human $CD34^+$ cells ($\sim$2 $\times$ 10^6 cells/mouse) were administered intravenously by retro-orbital injection to NSG female mice, which were given 1.5 Gy total body irradiation as a single dose, 3 hr before transplantation.

Animal study protocols were drafted in accordance with French and European legislation on animal experimentation and approved by the Genethon Institutional Animal Care and Use Committee.

Flow Cytometry

Erythroid differentiation was tested by flow cytometry after immunostaining for CD36 (Ab 656151; BD Pharmingen) and CD235a (Ab 563666; BD Pharmingen). Donor cell chimerism in total BM cells was tested by flow cytometry after immunostaining for human-specific CD45 (Ab A74763; Beckman Coulter). Human T, B, and myeloid phenotype in total BM and PB cells of transplanted mice was tested by flow cytometry after immunostaining for the human-specific CD45 and cell-specific antigens. For the analysis, cells were resuspended in PBS and incubated for 30 min at 4°C with antibodies for specific antigens or with isotype control following the manufacturer's instructions. Stained cells were washed and measured on BD LSR II flow cytometer (BD Biosciences). Analyses were performed with the FlowJo software v7.6.5.

VCN Evaluation

Genomic DNA from transduced cells expanded in liquid culture was extracted using the Wizard genomic DNA purification kit (Promega) following the manufacturer's instructions. Genomic DNA from hematopoietic colonies was extracted using Proteinase K (Thermo Fisher Scientific) lysis by adding 20 μL of lysis buffer (0.3 mM Tris HCl [pH 7.5]; 0.6 mM CaCl2; 1.5% glycerol; 0.675% Tween 20; and 0.3 mg/mL Proteinase K). Lysis was performed at 65°C for 30 min, 90°C for 10 min, and 4°C for 10 min. Finally, 30 μL of water was added to each sample. The VCN (vector genome per diploid genome [vg/dg]) was analyzed by duplex TaqMan qPCR with primers and probes annealing to the HIV *psi* sequence and a reference gene, the human *ALB* gene. Results were calculated based on a standard curve of a plasmid containing the two sequences. All PCR measurements were performed at least in duplicates in an ABI PRISM 7700 system (Thermo Fisher Scientific). All primer and probe sequences are listed in the Supplemental Materials and Methods.

Statistical Analysis

Statistical analyses in the results of the animal study were performed using GraphPad Prism version 7.0 for Windows. Results are reported as mean ± SE. Statistical differences between means were evaluated by Mann-Whitney or χ^2 test as appropriate. Differences were considered significant at p value <0.05.

Analysis of mRNA Expression

The βAS3 transgene expression was measured in transduced $CD34^+$ cells after erythroid differentiation in liquid bulk culture.

Total RNA was extracted from dry cell pellets using Roche RNeasy Kit (QIAGEN) following the manufacturer's instructions and treated with DNase (QIAGEN). Reverse transcription and qRT-PCR were performed with the LightCycler Multiplex RNA Virus Master Kit (Roche) according to the manufacturer's instructions. Duplex qPCRs with TaqMan probes were performed with primer pairs amplifying the AS3 transgene (but not the wild-type HBB sequence) and a normalizer gene (α-globin). Primers and probes are listed in the Supplemental Materials and Methods.

IVIM Assay

The IVIM assay was performed as previously described.[18,19,21] We transduced 1×10^5 cells and expanded them for 2 weeks, prior to limiting dilution replating with 100 cells per 96-well. VCN measurement was performed on day 4 after transduction. The MTT assay was used for quantification of cell proliferation in 96-well plates.[35] For determination of wells with proliferation, we used the highest absorbance value of the respective Mock plates, which were below a previously determine clonal proliferation level (5.61-fold the absorbance value of microscopically confirmed negative wells). Based on meta-data from positive and negative controls of previous assays, we defined values between the lower limit of detection (LOD = 1.05×10^{-4}) and a Q1 quantification limit (RF = 34.17×10^{-4}) as background.

Viral IS Recovery and Analysis

LTR vector-genome junctions were amplified by restriction-based or sonication-based LM-PCR. In the restriction-based method, 1 μg of genomic DNA was digested 6 hr at 37°C with the *Tru*91 restriction enzyme (Roche), purified by NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL) and ligated overnight to a TA-protruding double-stranded DNA linker by T4 DNA Ligase (New England Biolabs). Ligated DNA was purified, digested overnight at 37°C with *Sac*I (Roche), and purified again.

Multiple nested PCRs (10–20) were performed with specific primers annealing to the linker and the 3′ vector LTR, adapted to Illumina sequencing and containing a 4-bp sample-specific barcode for sample identification. Libraries were quantified on NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) and loaded on a 1% agarose gel for amplicon size selection. Amplicons ranging from 200 to 500 kb were manually extracted from the gel and purified by NucleoSpin Gel and PCR Clean-up kit. About 1 μg of the final libraries were subsequently processed with MiSeq Reagent Kit v3 (2 × 300 bp pair-end sequencing) following the manufacturer's instructions, bar-coded with a 6-bp sample-specific tag, and sequenced to saturation on an Illumina MySeq machine at IGA Technology Services (Udine, Italy). Raw reads resulting from Illumina paired-end sequencing were bioinformatically trimmed to recover the human genome sequences adjacent to the 3′ proviral LTR by the Skewer software (mismatch rate of 0.12, minimal length of 20 bp after trimming, and a minimal match length equal to the pre-trimming sequence) and mapped on the reference genome (human GRCh37.75/hg19) by the Bowtie2 software (>95% identity). The genomic coordinates of the first nucleotide in the host genome adjacent to the viral LTR were indicated as ISs. ISs originated by different reads mapping in the same genomic position were collapsed recovering the number of corresponding reads (read count). Unique ISs were annotated on the RefSeq gene database as intergenic or intragenic (exonic or intronic). Genes (defined by the most upstream transcription start site to the most downstream 3′ end) hosting at least one IS were identified as target genes. Genes with targeting frequency significantly higher than random, defined by 200 random sampling of virtual ISs corresponding to 5,330,124 and 5,812,911 *Tru*91 sites in the human and murine genome, respectively, were defined as over-targeted, after Bonferroni correction of the comparison p values. Genes with at least three ISs and a Bonferroni-corrected p value <0.001 were defined as over-targeted. Additional information about the bioinformatics processing of the sequencing reads can be found in the Supplemental Materials and Methods. Sequencing reads are available at Sequence Read Archive (SRA): PRJNA497990.

SUPPLEMENTAL INFORMATION

Supplemental Information includesone figure, six tables, and Supplemental Materials and Methods and can be found with this article online at https://doi.org/10.1016/j.omtm.2018.10.014.

AUTHOR CONTRIBUTIONS

Study Design: V.P., F.U., D.B.K., and F.M. Study Execution: V.P., F.U., S.C., G.C., R.P.H., B.C.F., S.M., and M.R. Data Analysis: V.P., F.U., S.C., G.C., R.P.H., B.C.F., M.R., A.S., D.B.K., and F.M. Manuscript Drafting: V.P., F.U., and F.M. Manuscript Reviewing, Discussion, and Finalization: V.P., F.U., M.R., A.S., D.B.K., and F.M.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

ACKNOWLEDGMENTS

This work was funded by the AFM-Téléthon, a French charitable organization. A.S. and M.R. were supported by the European Commission (FP7-Health-2010-CELL-PID, 261387), the Federal State of Lower Saxony (research project R2N), and DFG (SFB738, REBIRTH).

REFERENCES

1. Stamatoyannopoulos, G., ed. (2001). The Molecular Basis of Blood Diseases, Third Edition (W.B. Saunders Co.).
2. Madigan, C., and Malik, P. (2006). Pathophysiology and therapy for haemoglobinopathies. Part I: sickle cell disease. Expert Rev. Mol. Med. *8*, 1–23.
3. Platt, O.S. (2008). Hydroxyurea for the treatment of sickle cell anemia. N. Engl. J. Med. *358*, 1362–1369.
4. Locatelli, F., Kabbara, N., Ruggeri, A., Ghavamzadeh, A., Roberts, I., Li, C.K., Bernaudin, F., Vermylen, C., Dalle, J.H., Stein, J., et al.; Eurocord and European Blood and Marrow Transplantation (EBMT) Group (2013). Outcome of patients with hemoglobinopathies given either cord blood or bone marrow transplantation from an HLA-identical sibling. Blood *122*, 1072–1078.
5. Naldini, L. (2015). Gene therapy returns to centre stage. Nature *526*, 351–360.

6. Andreani, M., Testi, M., Gaziev, J., Condello, R., Bontadini, A., Tazzari, P.L., Ricci, F., De Felice, L., Agostini, F., Fraboni, D., et al. (2011). Quantitatively different red cell/nucleated cell chimerism in patients with long-term, persistent hematopoietic mixed chimerism after bone marrow transplantation for thalassemia major or sickle cell disease. Haematologica *96*, 128–133.

7. Walters, M.C., Patience, M., Leisenring, W., Rogers, Z.R., Aquino, V.M., Buchanan, G.R., Roberts, I.A., Yeager, A.M., Hsu, L., Adamkiewicz, T., et al. (2001). Stable mixed hematopoietic chimerism after bone marrow transplantation for sickle cell anemia. Biol. Blood Marrow Transplant *7*, 665–673.

8. Wu, C.J., Gladwin, M., Tisdale, J., Hsieh, M., Law, T., Biernacki, M., Rogers, S., Wang, X., Walters, M., Zahrieh, D., et al. (2007). Mixed haematopoietic chimerism for sickle cell disease prevents intravascular haemolysis. Br. J. Haematol. *139*, 504–507.

9. Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H., and Steinberg, M.H. (2011). Fetal hemoglobin in sickle cell anemia. Blood *118*, 19–27.

10. Steinberg, M.H., Chui, D.H., Dover, G.J., Sebastiani, P., and Alsultan, A. (2014). Fetal hemoglobin in sickle cell anemia: a glass half full? Blood *123*, 481–485.

11. Ferrari, G., Cavazzana, M., and Mavilio, F. (2017). Gene therapy approaches to hemoglobinopathies. Hematol. Oncol. Clin. North Am. *31*, 835–852.

12. Ribeil, J.A., Hacein-Bey-Abina, S., Payen, E., Magnani, A., Semeraro, M., Magrin, E., Caccavelli, L., Neven, B., Bourget, P., El Nemer, W., et al. (2017). Gene therapy in a patient with sickle cell disease. N. Engl. J. Med. *376*, 848–855.

13. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene therapy in patients with transfusion-dependent β-thalassemia. N. Engl. J. Med. *378*, 1479–1493.

14. Levasseur, D.N., Ryan, T.M., Reilly, M.P., McCune, S.L., Asakura, T., and Townes, T.M. (2004). A recombinant human hemoglobin with anti-sickling properties greater than fetal hemoglobin. J. Biol. Chem. *279*, 27518–27524.

15. Levasseur, D.N., Ryan, T.M., Pawlik, K.M., and Townes, T.M. (2003). Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. Blood *102*, 4312–4319.

16. Romero, Z., Urbinati, F., Geiger, S., Cooper, A.R., Wherley, J., Kaufman, M.L., Hollis, R.P., de Assin, R.R., Senadheera, S., Sahagian, A., et al. (2013). β-Globin gene transfer to human bone marrow for sickle cell disease. J. Clin. Invest. *123*, 3317–3330.

17. Urbinati, F., Hargrove, P.W., Geiger, S., Romero, Z., Wherley, J., Kaufman, M.L., Hollis, R.P., Chambers, C.B., Persons, D.A., Kohn, D.B., and Wilber, A. (2015). Potentially therapeutic levels of anti-sickling globin gene expression following lentivirus-mediated gene transfer in sickle cell disease bone marrow CD34+ cells. Exp. Hematol. *43*, 346–351.

18. Modlich, U., Bohne, J., Schmidt, M., von Kalle, C., Knöss, S., Schambach, A., and Baum, C. (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. Blood *108*, 2545–2553.

19. Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A., and Baum, C. (2008). Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol. Ther. *16*, 718–725.

20. Stein, S., Ott, M.G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Krämer, A., Schwäble, J., Glimm, H., et al. (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat. Med. *16*, 198–204.

21. Poletti, V., Charrier, S., Corre, G., Gjata, B., Vignaud, A., Zhang, F., Rothe, M., Schambach, A., Gaspar, H.B., Thrasher, A.J., and Mavilio, F. (2018). Preclinical development of a lentiviral vector for gene therapy of X-linked severe combined immunodeficiency. Mol. Ther. Methods Clin. Dev. *9*, 257–269.

22. Modell, B. (2008). Global Epidemiology of Hemoglobin Disorders and Derived Service Indicators (World Health Organization), pp. 417–496.

23. Kanter, J., Walters, M.C., Hsieh, M.M., Krishnamurti, L., Kwiatkowski, J., Kamble, R.T., von Kalle, C., Kuypers, F.A., Cavazzana, M., Leboulch, P., et al. (2016). Interim results from a phase 1/2 clinical study of lentiglobin gene therapy for severe sickle cell disease. Blood *128*, 1176.

24. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature *467*, 318–322.

25. Marktel, S., Cicalese, M.P., Giglio, F., Scaramuzza, S., Calbi, V., Casiraghi, M., Ciotti, F., Lidonnici, M.R., Rossi, C., Masera, N., et al. (2017). Gene therapy for beta thalassemia: preliminary results from the PHASE I/II Tiget-Bthal trial of autologous hematopoietic stem cells genetically modified with GLOBE lentiviral vector. Blood *130*, 355.

26. Weber, L., Poletti, V., Magrin, E., Antoniani, C., Martin, S., Bayard, C., Sadek, H., Felix, T., Meneghini, V., Antoniou, M.N., et al. (2018). An optimized lentiviral vector efficiently corrects the human sickle cell disease phenotype. Mol. Ther. Methods Clin. Dev. *10*, 268–280.

27. Miccio, A., Cesari, R., Lotti, F., Rossi, C., Sanvito, F., Ponzoni, M., Routledge, S.J., Chow, C.M., Antoniou, M.N., and Ferrari, G. (2008). In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of beta-thalassemia. Proc. Natl. Acad. Sci. USA *105*, 10547–10552.

28. Lagresle-Peyrou, C., Lefrère, F., Magrin, E., Ribeil, J.A., Romano, O., Weber, L., Magnani, A., Sadek, H., Plantier, C., Gabrion, A., et al. (2018). Plerixafor enables safe, rapid, efficient mobilization of hematopoietic stem cells in sickle cell disease patients after exchange transfusion. Haematologica *103*, 778–786.

29. Poletti, V., and Mavilio, F. (2017). Interactions between retroviruses and the host cell genome. Mol. Ther. Methods Clin. Dev. *8*, 31–41.

30. Biasco, L., Ambrosi, A., Pellin, D., Bartholomae, C., Brigida, I., Roncarolo, M.G., Di Serio, C., von Kalle, C., Schmidt, M., and Aiuti, A. (2011). Integration profile of retroviral vector in gene therapy treated patients is cell-specific according to gene expression and chromatin conformation of target cell. EMBO Mol. Med. *3*, 89–101.

31. Biasco, L., Baricordi, C., and Aiuti, A. (2012). Retroviral integrations in gene therapy trials. Mol. Ther. *20*, 709–716.

32. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M.P., et al. (2016). In vivo tracking of human hematopoiesis reveals patterns of clonal dynamics during early and steady-state reconstitution phases. Cell Stem Cell *19*, 107–119.

33. Moiani, A., Paleari, Y., Sartori, D., Mezzadra, R., Miccio, A., Cattoglio, C., Cocchiarella, F., Lidonnici, M.R., Ferrari, G., and Mavilio, F. (2012). Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. J. Clin. Invest. *122*, 1653–1666.

34. Cesana, D., Sgualdino, J., Rudilosso, L., Merella, S., Naldini, L., and Montini, E. (2012). Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. J. Clin. Invest. *122*, 1667–1676.

35. Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J. Immunol. Methods *65*, 55–63.